Portraits of micro-workers: The real people behind AI in France

Paola Tubaro, Antonio A. Casilli

▶ **To cite this version:**

Paola Tubaro, Antonio A. Casilli. Portraits of micro-workers: The real people behind AI in France. 2nd Crowdworking Symposium 2020, Research program "Digital Future", Universities of Paderborn and Bielefeld, Oct 2020, Paderborn, Germany. hal-02960775

HAL Id: hal-02960775
https://hal.science/hal-02960775

Submitted on 7 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Portraits of micro-workers

The real people behind AI in France

PAOLA TUBARO, National Center of Scientific Research (CNRS)
ANTONIO A. CASILLI, Telecom Paris

1 INTRODUCTION

Fueled by the simultaneous advent of "big" data and massive computational power, recent progresses in artificial intelligence (AI) have spread the belief that the current wave of automation may produce unprecedented socio-economic effects [Brynjolfsson and McAfee 2014; Frey and Osborne 2017]. In particular, popular "robots will steal our jobs" prophecies have revamped classical concerns [Ricardo 1821] that machines may supersede human labor.

Yet men and women contribute to producing AI solutions [Irani 2016], largely out of sight of the general public. *Micro-work* is a case in point. Data-related jobs are fragmented into myriad small tasks that can be performed remotely online, and specialized digital platforms such as Amazon Mechanical Turk, Microworkers and Clickworker allocate them to crowds of providers. Such tasks may consist in flagging inappropriate web content, labeling images, transcribing or translating bits of text, recording voice. Quick and repetitive, they are paid on a piecework basis, with rates as low as few cents. Their silent but essential contribution to the functioning of smart equipment, computer vision systems and virtual assistants, is what Gray and Suri [2017] call the "paradox of automation's last mile": if smart technologies destroy some traditional occupations, they also need the human-in-the-loop, so that the final outcomes result from a mix of machinery and people.

This paper addresses the two related questions of how micro-work underpins automation, and whose work it is – who are the real people behind today's AI, what motivates them to engage in this activity, and what conditions they face. It is shown that they include workers with underpriviledged backgrounds, downward life trajectories, or temporary difficulties due to phases of unemployment or care duties. The unregulated, unprotected setting of digital platform work is little suited to improve their chances, especially for the most vulnerable among them.

To support these claims, the paper draws on a mixed-methods study of micro-work in France, DiPLab ("Digital Platform Labor"), comprising a questionnaire distributed to about 1,000 micro-workers through Wirk[1], a prominent local micro-work platform, and 90 interviews with micro-workers, platform owners, business clients and other stakeholders. While the case of France is inherently interesting (a highly industrialized country and a pioneer in information technologies, yet less documented in the nascent literature on micro-work), it is also exemplary of more general behaviors and trends.

The study invites to reflect on how credible commitment to responsible, ethical AI requires considering micro-working conditions. Are digital platforms extending earning opportunities to the more fragile segments of society, those often excluded from standard labor markets (women, the elderly and/or disabled, minorities, people with lower skills)? Are they providing non-discriminatory environments where rewards depend only on performance? These are not intrinsically legal questions, but societal ones – albeit with ethical implications.

[1]The platform has just re-branded from its original identity as Foule Factory ("Foule" means "Crowd" in French), with specific focus on human-in-the-loop AI solutions.

Authors' addresses: Paola TubaroNational Center of Scientific Research (CNRS); Antonio A. CasilliTelecom Paris.

2 HOW MICROWORK FEEDS AI

Before discussing the place of micro-work in the AI supply chain, it must be distinguished from other types of digital platform work. It differs from *online freelancing*, which concerns more creative work (such as design and software development), involves qualified professionals, and entrusts them with full projects rather than single tasks. Micro-work also differs from *"gig" work* where platform-mediated services (such as urban transportation and food delivery) are geographically sticky [Graham et al. 2017] and therefore performed offline, even though coordination occurs online. What all these activities have in common is status of workers, who are not employed but contracted out for each task and paid based on their output, subject to algorithmic control performed by the platform.

Micro-work pushes to the extreme the "datafication" processes that underlie these activities, as its very essence is the production of data, the quality check of algorithmic outputs, and their replacement when they fail. [Casilli et al. 2019] call these three functions, respectively, "AI training", "AI validation" and "AI impersonation". In what follows, I illustrate them with the help of insight from our qualitative fieldwork.

2.1 AI training

A first way in which micro-work feeds AI is by generating so-called "training" datasets for machine-learning algorithms. For example, a task that several study participants did consisted in audio-recording themselves reading aloud short sentences. The goal of the client, producer of a virtual assistant, was to assemble a large set of diverse examples of voices, all pronouncing the same words, so that its vocal recognition algorithm could learn that they all meant the same despite differences in accent, background noise, etc. In other cases, the data are already available but cannot be used as such, and micro-work is needed to "annotate" them for better quality. For example, some respondents had to distinguish vegetables (tomatoes, carrots, etc.) in pictures of dishes. A trivial task for humans, recognizing a tomato is a challenge for a computer. Micro-workers help by drawing lines around tomatoes, circling them, or adding tags. With these annotations, computer-vision algorithms can compare different tomatoes and carrots across images and understand what they look like.

These routine tasks are relatively light to perform, and are often perceived as "silly". While some workers are just happy with the little pay they get for them, given the limited effort required, others find them more perplexing. One worker could not make sense of why she was told to “draw a square around a tomato” because “everyone knows what a tomato is”, and insisted that “I don’t know why”.

2.2 AI validation

If the first usage of micro-work in the machine-learning supply chain is at input level, to generate or enrich training datasets, another usage is at output level, to review the results. For example, one study participant checked the accuracy of speech recognition by a virtual assistant. She had to compare short audio recordings of what users said, to the transcriptions done automatically by the virtual assistant, to check if they were accurate, and if needed, to rectify. Other respondents reported checking the automated transcription of scanned receipts and invoices, adding corrections as appropriate. Only few of them understood that their work would be used by engineers to ensure their AI would not make the same mistakes in future; the great majority believed they were somehow helping some accountancy company.

2.3 AI impersonation

In some cases, micro-work does not support the processes of data generation or algorithmic quality assurance, but replaces (impersonates) them when they are not up to standards. This happens when humans outperform computers, either in terms of effectiveness or cost. The very idea that prompted Amazon to launch Mechanical Turk in 2006, was to integrate humans directly into software programming when relying on their contribution is more efficient than automating. According to [Irani 2015, p. 225], this platform was “born out of the failures of artificial intelligence to meet the needs of internet companies".

AI impersonation may become sinister when use of low-paid humans instead of algorithms is undisclosed. Sometimes under pressure from investors, start-ups sell alleged AI solutions that are in fact the product of micro-work – often outsourced to providers in French-speaking, low-income African countries. One interviewee, an entrepreneur, derided his competitors who had made Madagascar “the French AI leader”.

2.4 A global, long-run transformation of labor

The three cases just outlined hint that the need for micro-work is unlikely to be a temporary one. As more industrial sectors integrate AI-based solutions, the need for micro-workers to produce AI training datasets and to perform AI quality checks will remain high. As long as wealth and income disparities across countries maintain repositories of cheap human labor in the developing world, AI impersonation will remain a cost-effective option.

Yet micro-work remains in the background. Lack of transparency around human intervention can be misleading for customers of AI-based solutions, and detrimental to workers when they are incapable to give meaning to what they do. Micro-work is also a global phenomenon that, beyond the “computer will steal our jobs” rhetoric, creates competition between workers from different countries.

3 WHO ARE THE MICRO-WORKERS?

If micro-work is here to stay, it is important to understand what parts of society are likely to be most affected. Beyond micro-workers from low-income countries, who are the people who engage in this activity in an industrialized country like France? What follows characterizes them with the help of, primarily, questionnaire data, and secondarily, qualitative interviews.

3.1 Mostly women in active age

Women are over-represented in the surveyed French population of micro-workers active on the Wirk platform: they are 56.1% of all workers, against 51.6% women in the general French population as measured by Insee, the National Statistical Institute.

It is interesting to compare this result to other studies. A long-term demographic survey of Mechanical Turk observes about equal participation of men and women, although percentages differ greatly across countries [Difallah et al. 2018]. A study commissioned by the European Parliament that surveyed European micro-workers on 4 international platforms (Mechanical Turk, Clickworker, Crowdflower and Microworkers), reports 60% men [Forde et al. 2017]. A recent ILO research report finds a more skewed distribution, with women representing only one out of every three micro-workers, and one out of five in developing countries [Berg et al. 2018].

This suggests specialization: women and men are both active as micro-workers, but they differ in their choice of platforms. This may have to do with platforms’ characteristics: for example, fieldwork suggests that some French-speaking users avoid international platforms whose home page is only available in English.

The reasons of women’s micro-work become clearer by looking at their age. Women aged 25-44 are 66.9% in our sample, clearly an over-representation compared to the general French population (31.2%). This suggests a linkage between micro-work and family duties, confirmed by interviews. Many women respondents combine a part-time job with childcare, and do micro-tasks as an extra earning activity from home – in a hidden, unrecognized “triple working day”.

3.2 Highly educated workers, often unemployed

Highly educated people are over-represented among micro-workers: 66.5% have a bachelor’s degree or higher, in comparison to just 27.8% in the general French population. Similar tendencies were found in studies of other platforms [Berg et al. 2018].

The majority practice micro-work as a complement to a full- or part-time job which, for just over half of them, is a clerical occupation. Among those who do not work, 10.2% are students, 5.2% are pensioners, and 20.8% have no professional activity. The latter figure is of particular concern if compared to the unemployment rate of France, which according to Insee figures for 2017, is 8.9%. While the official statistics measure is more restrictive, mere definitional discrepancies would not suffice to explain the large gap between the two, which It largely mirrors a substantive fact: the unemployed are over-represented among micro-workers.

3.3 A lower-income population that needs extra earnings

Are micro-workers poorer than the general population? To answer this question, reference can be made to the French Observatory

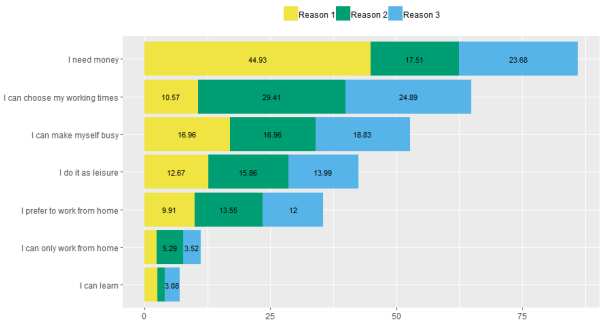

Fig. 1. Reasons for micro-work. Each participant was invited to choose three of them. Source: Author elaboration with DiPLab survey data, n = 908.

of Inequalities which defines the lower classes as the 30% of the population with the lowest incomes. Among the surveyed micro-workers, the lower classes are largely over-represented, with 51% meeting this criterion.

An alternative criterion is the poverty threshold, computed as 50% of the country's median income. 22.8% of the micro-workers under study live under the poverty threshold, so defined, compared with 8% of the general French population.

Against this background, online micro-tasks may be an attempt to cope. The survey asked respondents three reasons why they micro-work. The great majority say they need money – either as their first, second or third answer (Figure 1). In interviews, some said they actively look for better-paid tasks and feel frustrated if they turn out to be less rewarding than expected.

Yet earnings from micro-work are low: 22 euros a month on average, although the distribution is very skewed with a small number of people who make up to 2,000 euros a month, sometimes through multi-activity over different platforms.

4 CONCLUSIONS

Micro-work offers an opportunity for extra earnings to women with childcare duties who do not have a full-time job and do tasks primarily from home. It also attracts the unemployed and more generally, lower-income (albeit highly-educated) persons. However, these extra earnings are often tiny. The extent to which micro-work offers an opportunity for personal or professional development is doubtful, to the extent that tasks are mostly unqualified, that some micro-workers cannot give meaning to them, and that incentives to engage in more complex activities (such as signing up to foreign-language platforms) are scarce. While the study presented here focuses on one single country, France, its results are likely transposable to other settings.

Concluding, current reflections about ethical AI should fully integrate the issue of labor standards in the backstage of automation. Economic aspects and companies' cost considerations affect decisions and outcomes as much as techno-scientific progress. More importantly, the current development of AI is not independent from socio-economic inequalities inherited from the past, both between and within countries.

ACKNOWLEDGMENTS

Coordinated by Antonio A. Casilli and Paola Tubaro, the DiPLab ("Digital Platform Labor") research project is co-funded by Maison des Sciences de l'Homme Paris-Saclay; Force Ouvrière, a workers' union, through IRES (Social and Economic Research Institute); and France Stratégie, a service of the French Prime Minister. We also thank Wirk for logistical and financial support when fielding our survey, and Inria for complementary funding. For more information: http://diplab.eu/

Many thanks to current and former DiPLab team members for their inputs to this paper: Antonio A. Casilli, Maxime Besenval, Odile Chagny, Marion Coville, Clément Le Ludec, Touhfat Mouhtare, Lise Mounier, and Manisha Venkat. The usual disclaimer applies.

REFERENCES

J. Berg, M. Furrer, E. Harmon, U. Rani, and M.S. Silberman. 2018. *Digital labour platforms and the future of work: Towards decent work in the online world.* ILO Report.

E. Brynjolfsson and A. McAfee. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.* Norton.

A.A. Casilli, M. Coville, and P. Tubaro. 2019. All Work And No A.I.: How Human Labor Is A Lasting Feature Of Automation. *Working paper* (2019).

D. Difallah, E. Filatova, and P. Ipeirotis. 2018. Demographics and Dynamics of Mechanical Turk Workers. In *Proceedings of WSDM 2018: the Eleventh ACM International Conference on Web Search and Data Mining.* ACM, 135–143.

C. Forde, M. Stuart, S. Joyce, L. Oliver, D. Valizade, G. Alberti, K. Hardy, V. Trappmann, C. Umney, and C. Carson. 2017. *The Social Protection of Workers in the Collaborative Economy.* Report for the committee for employment and social affairs. European Parliament.

C.B. Frey and M. Osborne. 2017. The Future of Employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change* 114, C (2017), 254–280.

M. Graham, I. Hjorth, and V. Lehdonvirta. 2017. Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods. *Transfer: European Review of Labour and Research* 23, 2 (2017), 135–162.

M.L. Gray and S. Suri. 2017. The Humans Working Behind the AI Curtain. *Harvard Business Review* 9 (2017), 2–5.

L. Irani. 2015. Difference and dependence among digital workers: The case of Amazon Mechanical Turk. *South Atlantic Quarterly* 114 (2015).

L. Irani. 2016. *The Labor that Makes AI "Magic".* AINow Summit. White House and NYU.

D. Ricardo. 1821. *On the Principles of Political Economy and Taxation, 3rd edition.* John Murrey, Albemarle Street.